CO⋯ 2 −3 ⋯

HA⋯ .2.00

ML⋯ ⋯.0.50

42p

Western Management Science Institute
University of California ● *Los Angeles*

Working Paper No. 86

AN APPROACH TO STRICTLY CONCAVE PROGRAMMING WITH LINEAR CONSTRAINTS

by

Arthur M. Geoffrion

July 1965

1. INTRODUCTION

We shall give a finitely convergent procedure for solving the problem

(P)   Maximize $f(x)$ subject to $a_i x + b_i \geq 0$, $i = 1, \ldots, m$,

where f is a strictly concave and differentiable function of the n-vector x which assumes its unconstrained maximum. The a_i are n-vectors and the b_i are scalars.

No additional assumptions whatever, for example regarding the exclusion of degeneracy or the existence of a feasible solution of (P), are required. In fact the presence of degeneracy can only hasten convergence, and no perturbation, either real or virtual, is ever required. If (P) is infeasible, the procedure detects this fact routinely and terminates with a display of an infeasible subset of the constraints. Thus no feasible solution is needed to initiate the calculations.

An interesting aspect is that provision is made for making effective use of prior information which may be available regarding which constraints are likely to be binding at the optimum solution of (P) (under the tentative assumption that (P) is feasible).

The present method is not a "gradient" or "simplicial" or approximate method. It is aimed directly at constructing a solution of a certain version of the Kuhn-Tucker Conditions [7] for (P), which are necessary and sufficient for an optimal solution. It can be interpreted as solving a finitely convergent and "slowly changing" sequence of simpler maximization problems involving only linear equality constraints.

When $f(x)$ is quadratic the method specializes to a promising generalization of Theil and van de Panne's algorithm for quadratic

programming [8], with the added generality being that (a) certain

restrictive assumptions (regarding the feasibility of (P) and the

exclusion of degeneracy) are shown to be completely unnecessary, and (b)

provision is made for utilizing prior information regarding a given problem

so as to shorten the calculation time.

In order not to obscure the fundamental approach, in the next section

we derive the basic procedure and illustrate it by example under two

simplifying assumptions. These assumptions are removed in the following

section. Specialization to the quadratic case is the subject of the

penultimate section, and finally we discuss some computational considerations

and give a modification which permits linear equality constraints to be

handled efficiently.

Lengthy interruptions of the text are avoided by placing formal

statements and proofs of all theorems in the Appendix .

By way of preliminaries, we introduce two definitions. A subset S

of constraint indices is said to be consistent when the set $\{x: a_i x$

$+ b_i = 0, i \in S\}$ is not empty, and independent when the n-vectors a_i,

$i \in S$, are linearly independent. The empty set is considered to be both

consistent and independent. We denote by M the set of the first m posi-

tive integers, where m is the number of constraints of (P). The gradient

of f(x) is denoted by $\nabla f(x)$.

2. DEVELOPMENT UNDER TWO SIMPLIFYING ASSUMPTIONS

Throughout this paper, f(x) is assumed to be as stated in the introduction. For the sake of expository convenience, in this section we also temporarily assume (a) that (P) is feasible, which implies that (P) has a unique optimal solution x^* , and (b) that the set $B \equiv \{i \in M: a_i x^* + b_i = 0\}$ is independent.

Since our computational procedure is based on constructing a solution of the Kuhn-Tucker Conditions for (P) , we begin by recording without proof a version of Theorem 3 from [7], as specialized to (P) . This particular version was introduced in [6, Ch. 3].

<u>Theorem (Kuhn-Tucker)</u>:

Assume that f(x) is differentiable and concave. Then a trial solution x^o is an optimal solution of (P) if and only if for some subset $S^o \subseteq M$ of constraint indices there exist m real numbers (dual variables) u_i^o such that (x^o, u^o) satisfies the associated (Kuhn-Tucker) conditions:

$$(=S) \begin{cases} (KT-1) \quad \nabla f(x) + \sum_{i \in S} u_i a_i = 0 \\ (KT-2) \quad a_i x + b_i = 0, \ i \in S; \ u_i = 0, \ i \in M-S \end{cases}$$

$$(KT-3) \quad a_i x + b_i \geq 0, \ i \in M-S$$

$$(KT-4) \quad u_i \geq 0, \ i \in S .$$

The method suggested below for constructing a solution of the Kuhn-Tucker Conditions (KT-1,..., 4) involves solving for certain trial sets $S \subseteq M$, the equations (KT-1) and (KT-2) (together designated by the symbol (=S)) and then checking the inequations (KT-3) and (KT-4) to see if they are satisfied. If they are satisfied, the desired optimal solution of (P) is obviously at hand. This approach

suggests the following important definition. A set $S^o \subseteq M$ is said to be <u>valid</u> if and only if there exists a pair (x^o, u^o) which satisfies (KT-1, . . . , 4) associated with S^o . It is shown in Proposition 1 (see the Appendix for the proofs of this and subsequent propositions) that for any consistent and independent trial set $S \subseteq M$ the equations (=S) have a unique solution (x^S, u^S) , and it will be shown in section 2.1 below that there is at least one valid set, and that every valid set is consistent and independent. Thus we see that <u>the problem of solving (P) can be viewed as the problem of finding a</u> <u>valid set S^*</u> ; for then by the definition of validity (=S^*) yields as its unique solution an optimal solution x^* of (P) and the corresponding dual variables u^* . How one may go about finding a valid set is the main concern of the balance of this paper.

2.1 <u>Characterizing a Valid Set</u>

It is shown in Proposition 2 that there exists at least one triple (x^*, u^*, S^*) satisfying (KT-1, . . . , 4) , and that (x^*, u^*) is unique. It is clear from the nature of (KT-1, . . . , 4) that S^* can be taken as any set S satisfying $A \subseteq S \subseteq B$, where we make the definitions

$$A \equiv \{i \in M: \ u_i^* > 0\}$$
$$B \equiv \{i \in M: \ a_i x^* + b_i = 0\} .$$

In words, the set A is the collection of constraint indices corresponding to the constraints which are <u>active</u> at the optimum of (P) in the sense that their dual variables are strictly positive, and B is the set of indices corresponding to constraints which are <u>binding</u> at the optimum. From (KT-2) and (KT-4) , $A \subseteq B$. Conversely, it is clear that S^*

must satisfy $A \subseteq S \subseteq B$, or else (x^*, u^*, S) cannot satisfy

(KT-1, . . . , 4) .

From these and previous remarks we have the following characterizations

of validity: $S \subseteq M$ is valid if and only if $A \subseteq S \subseteq B$ if and only if

S is consistent and independent and (x^S, u^S) satisfies (KT-3) and

(KT-4) in addition to (=S) .

Unfortunately, neither A nor B are likely to be known a priori,

so that the identity of a valid set is not immediately available from

the first characterization. From the second characterization we do have,

however, (a) the useful necessary condition that a valid set must be

consistent and independent, which enables the search to be restricted to

such sets, and (b) a convenient test for validity to apply to such

candidates.

2.2 Determining the Order of Trials

The fact that there is but a finite number of subsets of M

immediately establishes the existence of a finite procedure for finding a

valid set--enumeration of consistent and independent sets. A non-repeating

sequence $\langle S^\upsilon \rangle$ of such trial sets terminates at the first trial for which

$(x^{S^\upsilon}, u^{S^\upsilon})$ satisfies (KT-3) and (KT-4) associated with S^υ , i.e.

at the first valid S^υ . When m is large, of course, undirected enumer-

ation is likely to be computationally impractical. We therefore develop

rules for directing the enumeration so as to keep the number of necessary

trials relatively small.

If S is not valid, then either $S - B \neq \emptyset$ or $A - S \neq \emptyset$, or both.

For obvious reasons the set $S - B$ will be called the excess of S ,

and $A - S$ will be called the deficiency of S . Clearly the smallest

change in S which would lead to a valid set would be to add its deficiency

and delete its excess. The number of indices in $(A -S)\cup(S-B)$ is
therefore a measure of the distance[1] $d(S)$ between S and the nearest
valid set. If S is consistent and independent, as well as invalid,
then by the second characterization of validity we have $a_i x^S + b_i < 0$
for some $i \epsilon$ M-S , or $u_i^S < 0$ for some $i \epsilon$ S , or both. An event of
the former type will be called a feasibility alarm, and an event of the
latter type an optimality alarm. A feasibility alarm from the deficiency
will be called a real alarm, as will an optimality alarm from the excess
of S .

An appealing conjecture is that the feasibility alarms coincide with
the deficiency, and that the optimality alarms coincide with the excess of
a consistent and independent but invalid set S . If this were true,
then by solving (=S) one could immediately determine the identity of a
valid set. Unfortunately this conjecture of perfect coincidence is not
true in general, as can easily be shown by counterexample (even for the
simple case $n = 2$ and $f(x)$ diagonal quadratic; see the example of
section 2.3). What can be shown (Proposition 3), however, is that at least
one alarm is real. This result suggests a procedure for determining the
order of trials by essentially heeding the alarms one at a time. By
heeding an alarm we mean adding to S a constraint which gives a feasibility
alarm, or deleting from S a constraint which gives an optimality alarm.
For convenience of exposition, we assume temporarily that S^0 is consistent

---

[1] The distance between a set $C \subseteq M$ and a set $D \subseteq M$ can be
defined as the number of elements in the symmetric difference set
$(C-D)\cup(D-C)$.

and independent whenever S is consistent and independent and i is a feasibility alarm due to S ; this assumption will then be removed.

Let S^o be an arbitrary consistent and independent initial trial set, and let it be invalid (i.e. let $d(S^o) \geq 1$). If S^o happens to yield exactly one alarm i_o , then by Proposition 3 it is real and therefore $d(S^o \pm i_o) = d(S^o) - 1$. Here we use the notation $S^o \pm i_o$ to mean $S^o \cup i_o$ or $S^o - i_o$ according as $i_o \notin S$ or $i_o \in S$ (i.e. according as i_o is a feasibility or an optimality alarm), in order to avoid having to distinguish between feasibility and optimality alarms. If S^o yields more than one alarm, then by Proposition 3 we have $d(S^o \pm i) = d(S^o) - 1$ for some $i \in T$, where T is the collection of alarms due to S^o . Therefore by trying each of the sets $S^o \pm i$, $i \in T$, we would obtain at least one set which is closer to the closest valid set (and, in fact, closer to every valid set). We call this the first generation of trials. If $d(S^o) = 1$, then a valid set is found during the first generation of trials; if $d(S^o) > 1$, a second generation of trials is necessary.

For each $i \in T$, define T_i to be the set of alarms generated by $S^o \pm i$. Since $d(S^o) > 1$, we have $T_i \neq \emptyset$ ($i \in T$) . The second generation of trials consists of the sets $S^o \pm i \pm j$ for all $i \in T$ and $j \in T_i$. The symbol $S^o \pm i \pm j$ means, of course, $(S^o \pm i) \cup j$ if $j \notin (S^o \pm i)$ and $(S^o \pm i) - j$ if $j \in (S^o \pm i)$. By Proposition 3, $d(S^o \pm i \pm j) = d(S^o) - 2$ for some $i \in T$ and $j \in T_i$. If $d(S^o) = 2$, then a valid set is found at this generation; if $d(S^o) > 2$, a third generation is necessary.

Continuing analagously, one constructs third and higher order generations as necessary. If at any trial a set is encountered which has

been tried before, it may, of course, be discarded.

It is clear that this search strategy terminates after exactly $d(S^o)$ generations of trials with the (unique) valid set which is closest to S^o , and that computational economies can be achieved due to the fact that changes in trial sets involve adding or deleting only one constraint at a time. A summary follows.

Rules for Ordering the Trials

0. Let S^o be consistent and independent, and let T be the set of alarms determined by S^o . If $T = \emptyset$, then S^o is valid and the procedure terminates; otherwise, a first generation of trials is necessary.

1. At the first generation of trials, for each $i \in T$ let T_i denote the set of alarms generated by $S^o \pm i$. If $T_i = \emptyset$ for some $i^* \in T$, then $S^o \pm i^*$ is valid and the procedure terminates; otherwise, a second generation of trials is necessary.

2. At the second generation of trials, for each $i \in T$ and $j \in T_i$, let T_{ij} be the set of alarms generated by $S^o \pm i \pm j$. If $T_{ij} = \emptyset$ for some $i^* \in T$, and $j^* \in T_i^*$, then $S^o \pm i^* \pm j^*$ is valid and the procedure terminates; otherwise, a third generation of trials is necessary.

3. Third and higher order generations of trials are constructed analagously.

.
.
.

(if at any trial a set is encountered which has been tried previously, it may be discarded).

It has been temporarily assumed that SUi_o is consistent and independent whenever S is and i_o is a feasibility alarm due to S . By the nature of the above rules, the effect of this assumption is to ensure that every trial set at every generation is consistent and independent, so that Proposition 3 can always be applied. This assumption may be dropped by appealing instead to Proposition 4, which asserts that SUi_o is independent when it is consistent, and that if SUi_o is inconsistent then the equations

$$\sum_{i \in S} z_i a_i + a_{i_o} = 0$$

have a unique solution z^S , $z_i^S < 0$ for at least one i in the excess of SUi_o , and $\{(SUi_o) - i\}$ is consistent and independent for any i such that $z_i^S < 0$. It follows that the above rules are applicable if we interpret "optimality alarms" as being defined for inconsistent sets SUi_o as the constraints corresponding to $z_i^S < 0$. The resulting procedure is given in Figure 1 (ignore for the time being the positive branch of step 2b, which is irrelevant when (P) is feasible).

Step 4 of Figure 1 can obviously oe omitted when an optimality alarm is oeing heeded, for the result can only be another consistent ana independent set. It can also be omitted when a feasibility alarm is being heeded. For example, when $f(x)$ is quadratic (see section 4) step 3 can go directly to step 1a and any inconsistency will be automatically detected during the calculations, thereby commanding transfer of control to step 1b. When a method is used for step 1a which does not routinely and quickly reveal inconsistency of the trial set, step 3 should go to step 1b when a feasibility alarm is being heeded--if the equations of step 1b have no solution (and efficient matrix partitioning methods [4; Appendix C of 6] should be used here) then control
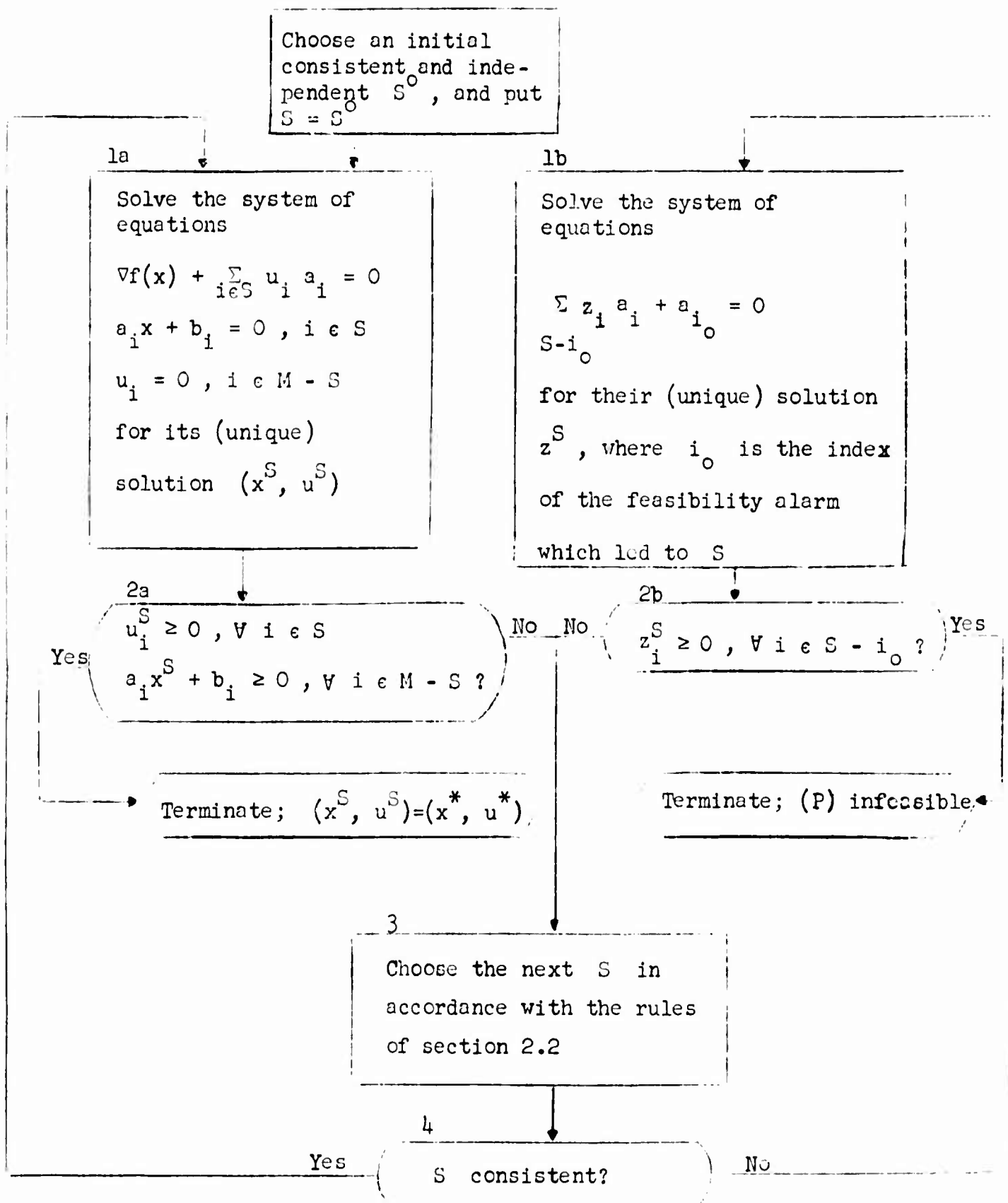
Choose an initial consistent and independent S^o, and put $S = S^o$

1a

Solve the system of equations

$$\nabla f(x) + \sum_{i \in S} u_i a_i = 0$$

$$a_i x + b_i = 0 , \; i \in S$$

$$u_i = 0 , \; i \in M - S$$

for its (unique)

solution (x^S, u^S)

1b

Solve the system of equations

$$\sum_{S - i_o} z_i a_i + a_{i_o} = 0$$

for their (unique) solution z^S , where i_o is the index of the feasibility alarm which led to S

2a

$u_i^S \geq 0 , \; \forall \; i \in S$

$a_i x^S + b_i \geq 0 , \; \forall \; i \in M - S$?

Yes   No   No

2b

$z_i^S \geq 0 , \; \forall \; i \in S - i_o$?

Yes

Terminate; $(x^S, u^S) = (x^*, u^*)$

Terminate; (P) infeasible

3

Choose the next S in accordance with the rules of section 2.2

4

Yes   S consistent?   No

Figure 1[*]

A Procedure for Concave Programming With Linear Constraints

[*]As noted in the text, step 4 can and should be incorporated (at little or no additional computational cost) into steps 1a or 1b.

must be transferred to step 1a.

Remark: If the initial trial set S^o can be chosen to be a subset of B , (e.g. $S^o = \emptyset$) then all optimality alarms can be ignored, as can any feasibility alarm which leads to an inconsistent set when heeded. This follows because any consistent and independent but invalid set with no excess gives at least one (feasibility) alarm from its deficiency, and heeding that alarm must obviously lead to another consistent and independent set. Similarly, if S^o can be chosen to contain A , then all feasibility alarms can be ignored. Thus certain types of prior information lead to considerable simplifications.

2.3 An Example

The example of Figure 2 is designed to illustrate the operation of the algorithm and show that not all alarms need be real. For convenience we present a graphical, rather than numerical, example in two dimensions (n=2) . We take $f(x)$ to be the ordinary distance from x to a fixed point x_o . By the Kuhn-Tucker Theorem, (=S) is a necessary and sufficient condition for a maximum of $f(x)$ over all x in the linear manifold $\{x: a_i x + b_i = 0, i \in S\}$. Hence (=S) yields as its solution the orthogonal projection x^S of x_o onto the manifold and the coefficients u_i^S of $-\nabla f(x^S)$ expressed in terms of the gradients a_i , $i \in S$. Loci of the four constraints are drawn and labeled with constraint indices; their gradients are also drawn in at selected points. The feasible region is hatched and x^* is marked by a heavy dot.

Let the initial consistent and independent trial set be arbitrarily chosen as, say, $S^o = \{3,4\}$. S^o is easily seen to yield a feasibility alarm for $i = 1$ and optimality alarms for $i = 3$ and 4 . Hence the
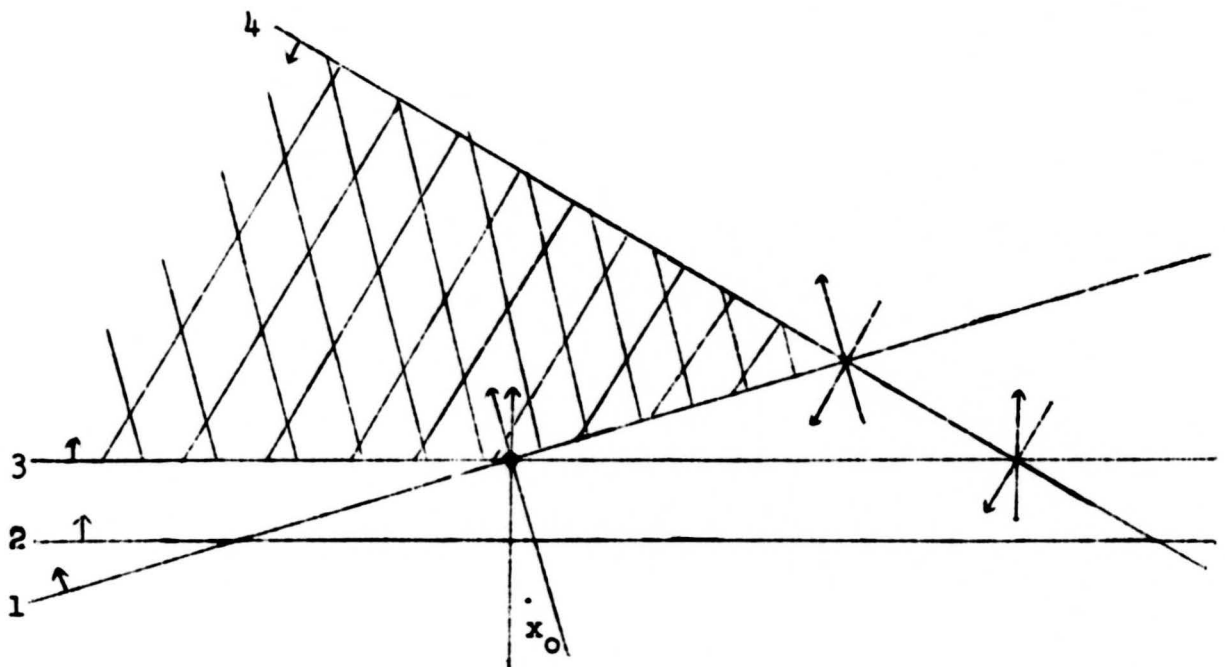
Figure 2

An Example

first generation trial sets are $(3,4,1)$, (4) , and (3) . Now (4) yields only one alarm, an optimality alarm fcr $i = 4$; (3) also yields one alarm, a feasibility alarm for $i = 1$; and $(3,4,1)$ is inconsistent-- solving $z_3 a_3 + z_4 a_4 + a_1 = 0$, one obtains optimality alarms for both $i = 3$ and 4 . Hence the second generation trial sets are $\emptyset$; $(3,1)$; $(4,1)$ and $(3,1)$. We find that $\emptyset$ yields feasibility alarms for $i = 1,2$, and 3 ; that $(4,1)$ yields optimality alarms for $i = 1$ and 4 ; and that $(1,3)$ yields no alarms at all. Hence $(1,3)$ is valid and the computations terminate. A diagrammatic summary of the trials is given in Figure 3. Of course, $A = B = (1,3)$ is obvious by inspection of Figure 2, so that a valid set has indeed been found. Note that $d((3,4)) = 2$, and that a valid set was indeed found in two generations.

Looking back over the calculations, we observe that $(3,4)$ gave rise to a optimality alarm for $i = 3$ which was not real; and that $\emptyset$ gave rise to a feasibility alarm for $i = 2$ which was not real. Hence not all alarms are real.
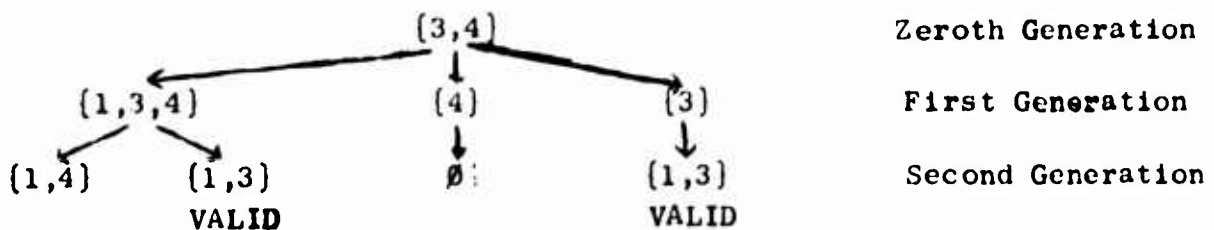

	$(3,4)$		Zeroth Generation
$(1,3,4)$	(4)	(3)	First Generation
$(1,4)$ $(1,3)$ VALID	$\emptyset$	$(1,3)$ VALID	Second Generation

Figure 3

A Summary of a Sequence of Trials

3. REMOVING THE SIMPLIFYING ASSUMPTIONS

It is now shown that the assumptions regarding the feasibility of (P) and the independence of B can be dropped without impairing the effectiveness of the procedure presented in the previous section.

3.1 Dropping the Assumption that B is Independent

If B is dependent, u^* need not be unique. Thus A need not be uniquely defined, and although the definition of a valid set remains as before we no longer have the characterization $A \subseteq S \subseteq B$ of validity. In place of A we define the collection A_j $(j = 1, \ldots, k)$ of all minimal valid sets, where a valid set is said to be minimal when no proper subset of itself is valid. It is shown in Proposition 5 that there is at least one minimal valid set and that each one is consistent, independent, a subset of B, and $u_i^{A_j} > 0$, $i \, \varepsilon \, A_j$. It is not difficult to see from the nature of the Kuhn-Tucker Conditions that a characterization of validity is $A_j \subseteq S \subseteq B$ for some $1 \le j \le k$. If S is consistent and independent, it is obvious that S is valid if and only if (x^S, u^S) satisfies (KT-3) and (KT-4).

The procedure of Figure 1 rests on Propositions 1 through 4. Propositions 1 and 4 and the first part of Proposition 2 do not require B to be independent. Proposition 5, as we have just shown, takes the place of the second part of Proposition 2. Proposition 6 is designed to take the place of Proposition 3. It follows that the procedure of Figure 1 still applies, and that at each generation of trials at least one trial set is one unit of distance closer to the collection of all valid sets than any trial set at the previous generation.

However, an exceptional case arises when a trial set S is not valid but nevertheless $x^S = x^*$, a possibility that can arise only when B is dependent (an example of this situation is given in the Appendix). In Proposition 6 it is shown that in this situation at least one optimality alarm is from $\bigcup_{j=1}^{k} (S - A_j)$, so that by heeding the alarms one at a time S will lead to at least one set which is closer to some A_j. Thus convergence is still assured.

3.2 Dropping the Assumption that (P) is Feasible

Assume that (P) has no feasible solution. We shall show that the procedure of Figure 1 is well-defined and terminates in the positive branch of step 2b at the first trial set encountered which corresponds to an infeasible subset of the constraints.

The initial trial set can be chosen to be the empty set, any singleton, or any other consistent and independent set as usual. By Proposition 1 the equations $(=S)$ have a unique solution (x^S, u^S) so long as S is consistent and independent, and clearly every such set leads to at least one feasibility alarm (otherwise (P) would be feasible). Furthermore, the first part of Proposition 4 still holds. It follows that the procedure of Figure 1 is well-defined and cannot terminate in the positive branch of step 2a. It remains to show that the procedure does terminate in the positive branch of step 2b.

A set $S \subseteq M$ of constraint indices is said to be <u>infeasible</u> when $\{x: a_i x + b_i \geq 0 , i \in S\}$ is empty, and is said to be <u>minimally</u> infeasible when no proper subset of itself is infeasible. Let $D_j (j = 1, \ldots, d)$ be the collection of distinct minimally infeasible sets. Then clearly $S \subseteq M$ is infeasible if and only if $D_j \subseteq S$ for some $1 \leq j \leq d$, and consistency implies feasibility but not conversely.

Every consistent and independent trial set S yields at least one feasibility alarm from every $D_j - S$, for otherwise some D_j would be feasible. Proposition 7 asserts that every inconsistent and feasible trial set which can arise during the procedure of Figure 1 yields at least one "optimality alarm" at step 2b from $\bigcup_{j=1}^{d} (S - D_j)$. Thus at each generation of trials at least one trial set is one unit of distance closer to the collection of all infeasible sets than any trial set at the previous generation, except for certain situations (associated with inconsistent and feasible trial sets) when one can assert only that at least one trial set is one unit of distance closer to at least one of the minimally infeasible sets. The occurrence of an inconsistent and infeasible trial set in a finite number of generations is thereby assured. A crude upper bound on the required number of generations is the distance from the initial trial set to the minimally infeasible set which is farthest from it. Proposition 7 also asserts that $z^S \geq 0$ for all infeasible sets which arise in the procedure of Figure 1. Thus termination in the positive branch of step 2b is assured.

<u>Remark</u>: The remark at the end of section 2.2 can be generalized to the case where it is not known whether or not (P) is feasible. Let the initial trial set be the empty set and ignore all optimality alarms as well as feasibility alarms which lead to an inconsistent set when heeded.

This strategy was seen to be justified when (P) is feasible. When (P) is infeasible, since trial sets can only grow as generations pass the situation will eventually be reached where a next generation cannot be defined because every feasibility alarm leads to an infeasible, and therefore inconsistent, set. This situation is the signal for termination when this strategy is used.

4. THE QUADRATIC CASE

Specialization to quadratic programming is natural and easy.

When $f(x) = \frac{1}{2}x^t Cx + c^t x$, where C is an n by n negative definite (symmetric) matrix and c is an n-vector, f meets the requirements stated in the introduction and (=S) is a system of linear equations with the unique solution (assuming that S is consistent and independent[2])

$$u_S^S = [A_S \, C^{-1} \, A_S^t]^{-1}(b_S - A_S \, C^{-1}c) \; ; \; u_i^S = 0 \; , \; i \in M\text{-}S$$
$$x^S = -C^{-1}(c + A_S^t \, u_S^S) \; ,$$

where t superscripts denote transpose, A_S is a matrix with rows a_i , $i \in S$, and u_S^S and b_S are similarly defined. A computer routine for implementing Figure 1 would invert C at the outset, along with $(A_{S^o} \, C^{-1} \, A_{S^o}^t)^{-1}$, and thereafter use efficient partitioning methods (see, e.g., [4] and Appendix C of [6]) for recomputing $(A_S \, C^{-1} \, A_S^t)^{-1}$ as S is altered by one constraint at a time. A similar strategy applies to the solution of $\sum_{i \in S-i_o} z_i a_i + a_{i_o} = 0$, which can be written (and this formula applies whether or not f is quadratic)

----

[2] If S were dependent, $[A_S \, C^{-1} \, A_S^t]^{-1}$ would fail to exist.

$$z^S = -a_{i_o} A^t_{S-i_o} (A_{S-i_o} A^t_{S-i_o})^{-1} .$$

Additional computational economies by partitioning can be obtained when some of the constraints are simple non-negativity constraints on the variables.

4.1 Relation to the Algorithm of Theil and van de Panne

The present algorithm specializes to Theil and van de Panne's algorithm [8] for quadratic programming under their assumptions[3]/ [8, p. 9], which can be phrased as follows: (a) (P) has at least two different feasible solutions, (b) C is negative definite, (c) $a_i x^S + b_i \neq 0$, i c M-S, for every consistent set S, and (d) the initial trial set is the empty set. Assumption (c) is strong and often difficult to verify. It can be shown to be equivalent to requiring for every consistent set S that if (x^S, u^S) satisfies (=S) then $u^S_i \neq 0$ for all i c S; in other words, degeneracy is completely excluded. Assumption (d) is not, of course, an assumption in the true sense of the word, but rather a convention--when it is followed, the remark at the end of section 2.2 applies and the procedure of Figure 1 is simplified.

The developments of this paper show that Theil and van de Panne's algorithm is not impaired if their assumption (c) is dropped, and that assumption (a) may also be dropped if it is understood that the infeasibility of (P) is distinguished by the occurrence of a generation at which at least one feasibility alarm occurs and every feasibility alarm leads to an

___

[3]/ Boot [1,2,3] has rederived the algorithm of Theil and van de Panne, but invokes similar assumptions (although he weakens assumption (c) somewhat).

inconsistent set. An important extension is obtained if assumption (d) is dropped when a priori information regarding B is available, for then S^o can very likely be chosen such that $d(S^o) < d(\emptyset)$ (hopefully $d(S^o) \ll d(\emptyset)$) and the calculation time can thereby be greatly reduced.

5. Discussion

It is evident that the algorithm of Figure 1 will be efficient when: (a) only a short time is required, on the average, to solve (=S') when (=S'') has been solved and S'' differs by only one index from S' ; and (b) S^o can be chosen so that $d(S^o)$ is reasonably small, as it often can be when a priori knowledge regarding A or B is available due to familiarity with the given problem or one quite similar to it (as in sensitivity analysis). Factor (a) determines the iteration time, and (b) the number of iterations.

When $f(x)$ is quadratic, it was observed that, due to linearity, (=S) can be solved very rapidly for a sequence of trial sets by making use of partitioning and bordering methods for matrix inversion and reinversion. In other cases, methods in numerical analysis for solving (partly nonlinear) systems of equations can be brought to bear, with simplifications arising for special classes of problems (such as linearly separable objective functions or non-negativity constraints). As an alternative to the direct solution of (=S) , one may wish to use the fact that x^S is characterized by the property that it maximizes $f(x)$ subject to $a_i x + b_i = 0$, $i \in S$. If x^S is found by a gradient-directed search or some other method, then u^S is easily found as $-\nabla f(x^S) A_S^t (A_S A_S^t)^{-1}$, where A_S is a matrix with rows a_i , $i \in S$ (partitioning methods can be used to keep track of the required inverse efficiently).

With regard to factor (b) it must be observed that when $d(S^0)$ is large not only might a very large number of trials be necessary, but also the amount of storage required for computer implementation of the rules of section 2.2 could become onerous. In this case other strategies for determining the sequence of trial sets should be considered. For example, if a large proportion of all alarms were real (as has been observed in numerous trial examples) then the following "Markov" strategy would prove to be an effective alternative to the one of section 2.2: heed one alarm at random. Storage is negligible, and although it is possible for this strategy to consume an infinite number of trials, the expected number of trials before convergence would be substantially smaller than for the previous strategy (for whatever consolation it may afford, we point out that it can be shown using the theory of Markov processes that the "Markov" strategy leads to convergence within a finite number of trials with probability one).

In closing, we observe that the algorithm can easily be modified to incorporate consistent and independent linear <u>equality</u> constraints[4] as well as inequality constraints in a more efficient manner than by rewriting them as pairs of inequalities (or using some similar device). It is easy to see that the indices of these constraints should be included in every trial set. This may be accomplished by including them in the initial trial set and then ignoring throughout the calculations any

---

[4] It can be shown that the hypothesis that f assumes its unconstrained maximum can be replaced in this case by the weaker hypothesis that f assumes its maximum over the linear manifold determined by the linear equality constraints.

optimality alarms which they happen to give. If the equality constraints are not independent, a maximal independent subset thereof should be used.

APPENDIX

In all of the propositions of this appendix, $f(x)$ is assumed to be differentiable, strictly concave, and to achieve its unconstrained maximum.

Lemma 1:

The maximum of $f(x)$ is uniquely achieved over every non-empty closed convex subset of n-dimensional Euclidean space E^n.

Proof (outline): Let X be a non-empty closed convex set in E^n. Put $Z = \{z: z \leq f(x)$ for some $x \in X\}$. It is easy to see that Z is non-empty, convex, and bounded from above; it can also be shown to be closed. The attainment of the maximum follows immediately, and is unique by the strict concavity of f.

Proposition 1:

If $S \subseteq M$ is consistent and independent, then the equations $(=S)$ have a unique solution (x^S, u^S).

Proof: Since S is consistent, the set $\{x: a_i x + b_i = 0, i \in S\}$ is non-empty as well as closed and convex. By Lemma 1 the maximum of f over this set is achieved uniquely. It is well-known that a solution (x^S, u^S) of $(=S)$ is a necessary and sufficient condition for this maximum. Hence x^S is unique. By the independence of S, u^S is also unique.

Proposition 2:

Assume that (P) is feasible. Then there exists a triple (x^*, u^*, S^*) satisfying $(KT-1,\ldots, 4)$, and x^* is unique. If, in addition, B is independent, then u^* is also unique.

Proof: By Lemma 1, (P) has a unique optimum x^* . The existence of u^* and $S^* \subseteq M$ such that (x^*, u^*, S^*) satisfies (KT-1,..., 4) is assured by the necessity of the Kuhn-Tucker Theorem; the uniqueness of x^* in such a triple follows from the sufficiency of the Kuhn-Tucker Theorem and the uniqueness of the optimal solution of (P) .

When B is independent, the uniqueness of u^* follows from (KT-1) upon using the uniqueness of x^* , the independence of B , and the fact that $u_i^* \neq 0$ only if $i \in B$ (which follows from the uniqueness of x^* and (KT-2)) .

Proposition 3:

Assume that (P) is feasible and that B is independent. Let S be consistent and independent but not valid. Then either $a_i x^S + b_i < 0$ for some $i \in A\text{-}S$, or $u_i^S < 0$ for some $i \in S\text{-}B$, or possibly both.

Proof: We proceed by contradiction. Suppose that (i) $a_i x^S + b_i \geq 0$, $i \in A\text{-}S$, and (ii) $u_i^S \geq 0$, $i \in S\text{-}B$.

Then from (ii), the Kuhn-Tucker Theorem, and the definition of (x^S, u^S) , it follows that x^S is an optimal solution of the problem

Maximize $f(x)$ subject to

$$a_i x + b_i = 0 , \quad i \in S \cap B$$
$$a_i x + b_i \geq 0 , \quad i \in S\text{-}B .$$

By (i), it follows that x^S is also optimal in the problem

Maximize f(x) subject to

(1) $a_i x + b_i = 0$, i ϵ S∩B

$a_i x + b_i \geq 0$, i ϵ (S-B) ∪ (A-S) .

Now by the Kuhn-Tucker Theorem, x^* is optimal in

(2) Maximize f(x) subject to $a_i x + b_i \geq 0$, i ϵ A .

Since the feasible region of (1) is included in that of (2), and since x^* is feasible in (1), x^* is also optimal in (1). But (1) has a unique optimal solution. Consequently, $x^S = x^*$.

To prove $u^S = u^*$, note that $x^S = x^*$ implies S ⊆ B and that (x^*, u^S) satisfies (=S) ; that (x^*, u^*) satisfies (=B) ; and apply the independence of B .

Thus (x^S, u^S) is identical with (x^*, u^*) , and therefore satisfies (KT-3) and (KT-4) as well as (KT-1) and (KT-2) . Consequently, we obtain the contradiction that S is valid.

Remark: Proposition 3 implies that when (P) is feasible and B is independent, heeding the alarms due to a consistent and independent set one at a time leads to at least one set which is one unit of distance closer to the closest valid set--and, in fact, one unit of distance closer to every valid set.

Lemma 2:

Let I_1 , I_2 , and I_3 be disjoint subsets of M . Exactly one of the following systems has a solution:

(i) $a_i x + b_i = 0$, i ϵ I_1

$a_i x + b_i \leq 0$, i ϵ I_2

$a_i x + b_i \geq 0$, i ϵ I_3

(ii) $\sum_{I_1 \cup I_2 \cup I_3} y_i \, a_i = 0$

$\sum_{I_1 \cup I_2 \cup I_3} y_i \, b_i + 1 = 0$

$y_i \leq 0 \ , \ i \in I_2$

$y_i \geq 0 \ , \ i \in I_3$

Proof: The proof of this lemma follows from the dualty theorem of linear programming [5].

System (i) has a solution iff the linear programming problem

(1) $\underset{x}{\text{Minimize}}$ ox subject to system (i)

is feasible. System (ii) has a solution iff the linear programming problem

$\underset{y}{\text{Maximize}} \ \sum_{I_1 \cup I_2 \cup I_3} y_i \, (-b_i)$ subject to

(2) $\sum_{I_1 \cup I_2 \cup I_3} y_i \, a_i = 0$

$y_i \begin{cases} \leq 0 \ , \ i \in I_2 \\ \geq 0 \ , \ i \in I_3 \end{cases}$

has an unbounded optimal value; for if y^o satisfies (ii) then θy^o is feasible in (2) for all real $\theta \geq 0$ and $\sum_{I_1 \cup I_2 \cup I_3} \theta y_i^o \, (-b_i) \equiv \theta \to \infty$ as $\theta \to \infty$, and conversely if the value of (2) can be made arbitrarily large then by the homogeneity of the constraints it can be made equal to unity. Observing that (1) and (2) are dual linear programming problems and that (2) is feasible (e.g. put $y = 0$), by the dualty theorem we have that (1) is infeasible iff (2) has an unbounded optimal value. This completes the proof.

Proposition 4:

Let $S \subseteq M$ be consistent and independent, and let $a_{i_o} x^S + b_{i_o} < 0$. Then $S \cup i_o$ is inconsistent if and only if it is dependent, and if it is dependent then

(1) $$\sum_{i \in S} z_i a_i + a_{i_o} = 0$$

has a unique solution z^S and $\{S \cup i_o - i\}$ is consistent and independent for any $i \in S$ such that $z_i^S < 0$. If, in addition, (P) is feasible, then inconsistency of $S \cup i_o$ implies that the excess of S is non-empty and that $z_i^S < 0$ for some i therein.

Proof: Since S is independent, $S \cup i_o$ is dependent iff (1) has a solution. By Lemma 2, $S \cup i_o$ is inconsistent iff

(2)
$$\sum_{i \in S} y_i a_i + y_{i_o} a_{i_o} = 0$$
$$\sum_{i \in S} y_i b_i + y_{i_o} b_{i_o} = -1$$

has a solution. Thus to show that $S \cup i_o$ is inconsistent iff it is dependent it is equivalent to show that (1) has a solution iff (2) does. To establish that (1) has a solution when (2) does, let y' be a solution of (2) and note that it follows from the definition of x^S and Lemma 2 that $y_{i_o}' > 0$; hence $z_i^S = (y_i'/y_{i_o}')$, $i \in S$, solves (1). To establish the converse, we shall show that $y' = \theta z_i'$, $i \in S$, and $y_{i_o}' = \theta$ solves (2) for some real $\theta > 0$ if z' solves (1). Suppose the contrary. Then it must be that

(3) $$\sum_{i \in S} z_i' b_i + b_{i_o} \geq 0,$$

and upon multiplying (1) by x^S and adding the result to (3) one would obtain

$$\sum_{i\in S} z_i' (a_i x^S + b_i) + (a_{i_o} x^S + b_{i_o}) \geq 0 \ ,$$

which would lead to a contradiction of the definition of x^S . The proof

that $S \cup i_o$ is inconsistent iff it is dependent is complete.

When $S \cup i_o$ is dependent, (1) must have a unique solution z^S by

the independence of S . To prove that $\{S \cup i_o - i\}$ is consistent and

independent for any i such that $z_i^S < 0$, it suffices to prove this

statement for any i such that $z_i^S \neq 0$. Independence is an easy con-

sequence of the independence of S . To show consistency, by Lemma 2

it is equivalent to show that the existence of $\tilde{y}$ such that

$\sum_{i\in S-i_*} \tilde{y}_i \, a_i + \tilde{y}_{i_o} \, a_{i_o} = 0$ and $\sum_{i\in S-i_*} \tilde{y}_i \, b_i + \tilde{y}_{i_o} \, b_{i_o} + 1 = 0$, where

$z_{i_*}^S \neq 0$, leads to a contradiction. The contradiction is obtained as

follows. Assume that $\tilde{y}_{i_o} \neq 0$, for otherwise by Lemma 2 $S - i_*$

would be inconsistent, which would contradict the consistency of S .

Then $\sum_{i\in S-i_*} (\tilde{y}_i / \tilde{y}_{i_o}) \, a_i + a_{i_o} = 0$. But by the uniqueness of the solution

of (1), this is impossible when $z_{i_*}^S \neq 0$.

Finally we prove that if (P) is feasible and $S \cup i_o$ is inconsistent,

then S has a non-empty excess and $z_i^S < 0$ for at least one i therein.

Suppose to the contrary that $z_i^S \geq 0$, $i \in S-B$, where we do not rule out

the possibility that $S-B = \emptyset$. Then for some $\theta^o > 0$ one would have

$$\sum_{i\in S} \theta^o \, z_i^S \, a_i + \theta^o \, a_{i_o} = 0$$

$$\sum_{i\in S} \theta^o \, z_i^S \, b_i + \theta^o \, b_{i_o} = -1 \ ,$$

where $\theta^o \, z_i^S \geq 0$, $i \in S-B$, and by Lemma 2 the system

$$a_i x + b_i = 0 , \ i \in S \cap B$$
$$a_i x + b_i \geq 0 , \ i \in \{S-B\} \cup i_o$$

would have no solution, thus contradicting the fact that

$$a_i x^* + b_i = 0 \ , \ i \ \epsilon \ B$$
$$a_i x^* + b_i > 0 \ , \ M\text{-}B \ .$$

This completes the proof.

Lemma 3:

Let d be a given n-vector. If $\sum_{i \epsilon S} u_i \ a_i = d$ has a non-negative solution, then it has a non-negative solution u^1 such that $\{i \ \epsilon \ S: \ u_i^1 > 0\}$ is independent.

Proof: See [5, p. 50].

Proposition 5:

Assume that (P) is feasible. Then the collection $A_j (j = 1, \ldots, k)$ of all minimal valid sets for (P) is not empty, and each A_j is consistent, independent, a subset of B, and $u_i^{A_j} > 0$ for each $i \ \epsilon \ A_j$. Furthermore, an arbitrary set S^o is valid if and only if $A_{jo} \subseteq S^o \subseteq B$ for some $1 \leq jo \leq k$.

Proof: By Proposition 2, there is at least one triple (x^*, u^*, S^*) which satisfies (KT-1,..., 4) . If S^* is not minimal, then from the nature of (KT-1,..., 4) and Lemma 3 it is clear that one may delete constraints from S^* and construct new couples (x^*, u) satisfying the associated (KT-1,..., 4) until a minimal valid set is obtained. Thus there is at least one minimal valid set.

Let A_j be any minimal valid set. By the definition of validity, there exists u^{A_j} (at this point we cannot assert the uniqueness of u^{A_j})

such that (u^{A_j}, x^*) satisfies (KT-1,..., 4) associated with A_j . It follows immediately that $A_j \subseteq B$, and is therefore consistent. From the nature of (KT-1,..., 4) , $u_i^{A_j} > 0$ for all $i \in A_j$; otherwise a proper subset of A_j would be valid. The independence of A_j follows by similar reasoning with the aid of Lemma 3.

The criterion for the validity of an arbitrary subset of constraint indices follows straightforwardly by arguments similar to those used above.

Proposition 6:

Assume that (P) is feasible. Let $S \subseteq M$ be consistent and independent but __not__ valid. If $x^S \neq x^*$, then for each minimal valid set A_j we have either $a_i x^S + b_i < 0$ for some $i \in A_j - S$, or $u_i^S < 0$ for some $i \in S-B$, or possibly both. If, on the other hand, $x^S = x^*$, then $\overset{k}{\underset{j=1}{\cup}} (S-A_j) \neq \emptyset$ and $u_i^S < 0$ for some i therein.

__Proof__: Assume $x^S \neq x^*$. We proceed by contradiction. Suppose for some $1 \leq j_0 \leq k$ that (i) $a_i x^S + b_i \geq 0$, $i \in A_{j_0} - S$, and (ii) $u_i^S \geq 0$, $i \in S-B$. Then by arguing exactly as in the proof of Proposition 3 with A_{j_0} in place of A , one may derive the contradiction that $x^S = x^*$.

Now we assume $x^S = x^*$. Observe that $S \subseteq B$, and therefore since S is not valid we have $A_j \not\subseteq S$, $j = 1,..., k$. Clearly $S - A_j \neq \emptyset$, $j = 1,..., k$, for otherwise S would have to be a proper subset of some A_j , which would contradict the uniqueness of the positive representation of $-\nabla f(x^*)$ in terms of the a_i , $i \in A_j$. A fortiori, $\overset{k}{\underset{j=1}{\cup}}(S - A_j) \neq \emptyset$. It remains to show that $u_i^S < 0$ for some i therein. A useful identity is

$$\overset{k}{\underset{j=1}{\cup}}(S - A_j) \equiv S - S \overset{k}{\underset{j=1}{\cap}} A_j ,$$

and for simplicity we write I for $S_j\bigcap_{j=1}^{k} A_j$.

We may assume: (a) that $I \neq \emptyset$, for otherwise $S - I \equiv S$ and hence the desired conclusion (which becomes $u_i^S < 0$ for some $i \in S$) must obtain or else (x^S, u^S) would satisfy (KT-4) as well as (KT-1, 2, 3) and S would be valid; and (b) that $u_i^S < 0$ for some $i \in I$, for otherwise the desired conclusion must obtain or else $u_i^S \geq 0$ for all $i \in S$ and S would again be valid. We shall prove that $u_i^S < 0$ for some $i \in S - I$ by supposing the contrary and finding $i_o \in I$ such that $S_j\bigcup_{j=1}^{k} A_j - i_o$ is valid, thereby contradicting the fact that every valid set contains at least one A_j . To find such an i_o , it suffices to find $i_o \in I$ and $\tilde{u}$ such that

(1) $\qquad \nabla f(x^*) + \sum_{i \in \{S_j\bigcup_{j=1}^{k} A_j - i_o\}} \tilde{u}_i\, a_i = 0 \qquad$ and $\qquad \tilde{u}_i \geq 0 , \; i \in \{S_j\bigcup_{j=1}^{k} A_j - i_o\}$;

for then $(x^*, \tilde{u})$ satisfies (KT-1,..., 4) associated with $\{S_j\bigcup_{j=1}^{k} A_j - i_o\}$. We now proceed to find such an i_o and $\tilde{u}$.

Now for each minimal valid set, we have $u_i^{A_j} > 0$, $i \in A_j$, and u^{A_j} satisfies

$$\nabla f(x^*) + \sum_{i \in A_j} u_i^{A_j} a_i = 0 .$$

For convenience, we write u_{ij} for $u_i^{A_j}$. Partitioning A_j into I and $A_j - I$, adding all k such equations and dividing by k , one obtains

(2) $\qquad \nabla f(x^*) + \sum_{i \in I} \sum_{j=1}^{k} (u_{ij}/k)\, a_i + \sum_{j=1}^{k} \sum_{i \in A_j - I} (u_{ij}/k)\, a_i = 0 .$

From assumption (b) and our supposition we also have $u_i^S \geq 0$, $i \in S - I$, and $u_i^S < 0$ for some $i \in I$, and u^S satisfies

$$(3) \qquad \nabla f(x^*) + \sum_{i \in S-I} u_i^S a_i + \sum_{i \in I} u_i^S a_i = 0 \; .$$

Multiplying (2) by $(1-t)$ and (3) by t and adding, where t is a real number, we obtain

$$(4) \qquad \nabla f(x^*) + \sum_{i \in I} [t\, u_i^S + (1-t) \sum_{j=1}^{k}(u_{ij}/k)]\, a_i + t \sum_{i \in S-I} u_i^S a_i +$$

$$(1-t) \sum_{j=1}^{k} \sum_{i \in A_j -I} (u_{ij}/k)\, a_i = 0 \; .$$

Put

$$t^o = \underset{\{i \in I: \; u_i^S < 0\}}{\mathrm{Min}} \left\{ \sum_{j=1}^{k}(u_{ij}/k) \Big/ (\sum_{j=1}^{k}(u_{ij}/k) - u_i^S) \right\} \; .$$

By construction, clearly $0 < t^o < 1$ and $[t^o u_i^S + (1-t^o) \sum_{j=1}^{k}(u_{ij}/k)] \geq 0$ for all $i \in I$ with strict equality holding for at least one i , say i_o . Upon making the appropriate identifications of $\tilde{u}$ in (4) with $t = t^o$, it follows that (1) holds. This completes the proof.

Proposition 7:

Assume that (P) is infeasible. Let $S \subseteq M$ be consistent and independent, and let $S \cup i_o$ be inconsistent, where $a_{i_o} x^S + b_{i_o} < 0$. Then

$$(1) \qquad \sum_{i \in S} z_i a_i + a_{i_o} = 0$$

has a unique solution z^S , and $z^S \geq 0$ if and only if $S \cup i_o$ is in-feasible. Furthermore, if $S \cup i_o$ is feasible then $\bigcup_{j=1}^{d} \{(S \cup i_o) - D_j\} \neq \emptyset$ and $z_i^S < 0$ for some i therein.

Proof: The existence of the unique solution z^S of (1) is known from

Proposition 4, and the assertion that $z^S \geq 0$ iff SUi_o is infeasible is a consequence of Lemma 2 and the relationship established in the proof of Proposition 4 between the solutions of (1) and of

(2a) $$\sum_{i \in SUi_o} y_i a_i = 0$$

(2b) $$\sum_{i \in SUi_o} y_i b_i + 1 = 0 .$$

Assume that SUi_o is feasible. This, with the inconsistency of SUi_o and the definition of x^S, implies by Lemma 2 that (2) has a solution y' such that $y'_{i_o} > 0$ and $y'_i < 0$ for some $i \in S$. First we show that $\{(SUi_o) - D_j\} \neq \emptyset$, $j = 1, \ldots, d$. Suppose the contrary. Then $SUi_o \subseteq D_{j_o}$ for some $1 \leq j_o \leq d$, and in fact $SUi_o \subset D_{j_o}$ since SUi_o is feasible. Since D_{j_o} is minimally infeasible, there exist real numbers $w_i > 0$, $i \in D_{j_o}$, such that

(3a) $$\sum_{i \in \{D_{j_o} - SUi_o\}} w_i a_i + \sum_{i \in SUi_o} w_i a_i = 0$$

(3b) $$\sum_{i \in \{D_{j_o} - SUi_o\}} w_i a_i + \sum_{i \in SUi_o} w_i b_i + 1 = 0 .$$

Letting $y = y'$ in (2) and taking a convex combination of (2a) with (3a), and of (2b) with (3b), one obtains for any real t

(4a) $$\sum_{i \in \{D_{j_o} - SUi_o\}} (1-t) w_i a_i + \sum_{i \in SUi_o} [(1-t) w_i + t y'_i] a_i = 0$$

(4b) $$\sum_{i \in \{D_{j_o} - SUi_o\}} (1-t) w_i b_i + \sum_{i \in SUi_o} [(1-t) w_i + t y'_i] b_i + 1 = 0 .$$

Since $y'_i < 0$ for some $i \in S$, $y'_{i_o} > 0$, and $w_i > 0$, $i \in D_{j_o}$, we can choose an appropriate $0 < t^o < 1$ to make $[(1-t^o) w_i + t^o y'_i] \geq 0$, $i \in SUi_o$, with strict equality holding for some $i \in S$. Hence a proper

subset of D_{j_o} would, by Lemma 2, be infeasible--thus violating the minimal infeasibility of D_{j_o} . Hence our supposition must be wrong and a fortiori $\overset{d}{\underset{j=1}{\cup}}\{(S \cup i_o) - D_j\} \neq \emptyset$, $j = 1,\dots, d$. It remains to show that $z_i^S < 0$ for some i therein. A useful identity is

$$\overset{d}{\underset{j=1}{\cup}}\{(S \cup i_o) - D_j\} \equiv S \cup i_o - (S \cup i_o) \overset{d}{\underset{j=1}{\cap}} D_j \ ,$$

and for simplicity we write I for $(S \cup i_o) \overset{d}{\underset{j=1}{\cap}} D_j$.

We may assume (a) that $I \neq \emptyset$, for otherwise $\overset{d}{\underset{j=1}{\cup}}\{(S \cup i_o) - D_j\} \equiv S \cup i_o$ and the desired conclusion follows from the known fact that $z_i^S < 0$ for some $i \in S$, and (b) that $z_i^S < 0$ for some $i \in I$, for otherwise the desired conclusion again follows from the known fact that $z_i^S < 0$ for some $i \in S$. Using these assumptions, we shall prove the desired assertion by supposing the contrary and finding $i_* \in I$ such that $(S \cup i_o) \overset{d}{\underset{j=1}{\cup}} D_j - i_*$ is infeasible, thereby contradicting the fact that every infeasible set contains at least one D_j . To find such an i_* , it suffices to find $i_* \in I$ and $\tilde{y} \geq 0$ such that

$$\underset{i \in \{(S \cup i_o) \overset{d}{\underset{j=1}{\cup}} D_j - i_*\}}{\sum} \tilde{y}_i a_i = 0$$

$$\underset{i \in \{(S \cup i_o) \overset{d}{\underset{j=1}{\cup}} D_j - i_*\}}{\sum} \tilde{y}_i b_i + 1 = 0 \ ,$$

for then by Lemma 2 $\{(S \cup i_o) \overset{d}{\underset{j=1}{\cup}} D_j - i_*\}$ would be infeasible. We now proceed to find such an i_* and $\tilde{y}$ under the supposition that $z_i^S \geq 0$, $i \in S \cup i_o - I$.

Since each D_j is minimally infeasible, there exist real numbers $w_{ij} > 0$, $i \in D_j$, $j = 1,\dots, d$, such that

(5a) $\sum\limits_{i \in D_j} w_{ij}\, a_i = 0$ and

(5b) $\sum\limits_{i \in D_j} w_{ij}\, b_i + 1 = 0$.

Partitioning D_j into I and $D_j - I$, adding all d equations of type (5a) and dividing by d , one obtains

(6a) $\sum\limits_{i \in I} \sum\limits_{j=1}^{d} (w_{ij}/d)\, a_i + \sum\limits_{j=1}^{d} \sum\limits_{i \in D_j - I} (w_{ij}/d)\, a_i = 0$.

Similarly, the equations of type (5b) yield

(6b) $\sum\limits_{i \in I} \sum\limits_{j=1}^{d} (w_{ij}/d)\, b_i + \sum\limits_{j=1}^{d} \sum\limits_{i \in D_j - I} (w_{ij}/d)\, b_i + 1 = 0$.

Putting $y = y'$ in (2) and adding t times (2a) to $(1-t)$ times (6a) yields

(7a) $\sum\limits_{i \in I} [t\, y_i' + (1-t) \sum\limits_{j=1}^{d} (w_{ij}/d)]\, a_i + \sum\limits_{i \in S \cup i_o - I} t\, y_i'$

$+ \sum\limits_{j=1}^{d} \sum\limits_{i \in D_j - I} (1-t)\, (w_{ij}/d)\, a_i = 0$.

Similarly, (2b) and (6b) yield

(7b) $\sum\limits_{i \in I} [t\, y_i' + (1-t) \sum\limits_{j=1}^{d} (w_{ij}/d)]\, b_i + \sum\limits_{i \in S \cup i_o - I} t\, y_i'$

$+ \sum\limits_{j=1}^{d} \sum\limits_{i \in D_j - I} (1-t)\, (w_{ij}/d)\, b_i + 1 = 0$.

Since the signs of z_i^S and y_i' agree for $i \in S$, by assumptions (a) and (b) and our supposition we can choose an appropriate $0 < t^o < 1$ to make $[t^o\, y_i' + (1-t^o) \sum\limits_{j=1}^{d} (w_{ij}/d)] \geq 0$, $i \in I$, with strict equality

holding for some i_* therein. Upon making the appropriate identifications in (7a) and (7b), the desired i_* and $\tilde{y}$ are at hand. This completes the proof.

Example

The example of Figure A-1 is designed to show that $x^S = x^*$ is possible when S is consistent, independent and invalid and B is dependent. The same conventions are followed here as in the example of section 2.3 of the text.
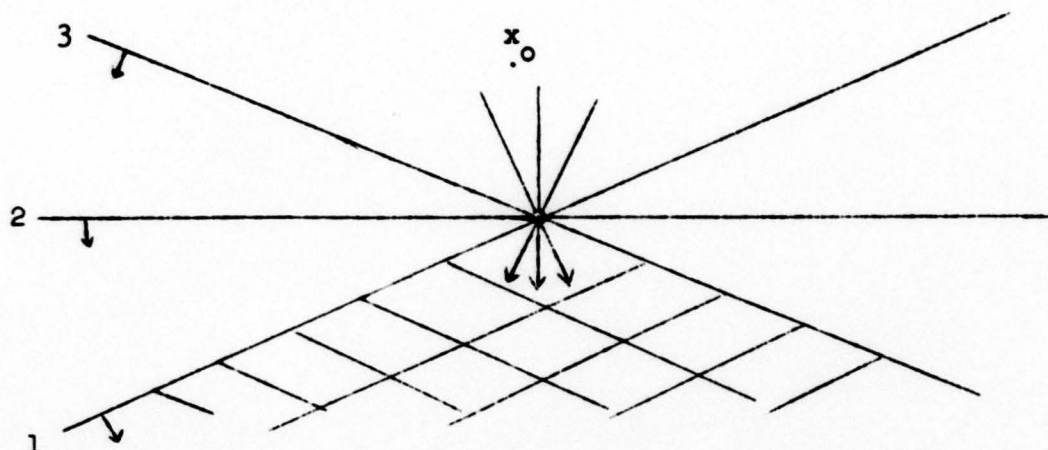

Figure A-1

Clearly $B = \{1,2,3\}$, and we may take the minimal valid sets as $A_1 = \{1,2\}$ and $A_2 = \{1,3\}$ (k = 2) . Thus the consistent and independent trial set $\{2,3\}$ is not valid. Nevertheless, $x^{\{2,3\}} = x^*$.

In accordance with the pertinent assertion of Proposition 6, $S = \{2,3\}$ yields an optimality alarm in $\{S - A_1\} \cup \{S - A_2\} \equiv \{2,3\}$, namely for i = 3 . Heeding it leads to the trial set $\{2\}$, which yields a feasibility alarm for i = 1 ; heeding this feasibility alarm leads to $\{2,1\}$, which is valid (i.e. it yields no alarms).

REFERENCES

1. Boot, J. C. G., "Notes on Quadratic Programming: The Kuhn-Tucker and Theil-van de Panne Conditions, Degeneracy, and Equality Constraints," *Management Science*, Volume 8, October 1961, 85-98.

2. Boot, J. C. G., "On Trivial and Binding Constraints in Programming Problems," *Management Science*, Volume 8, July 1962, 419-441.

3. Boot, J. C. G., "Binding Constraint Procedures of Quadratic Programming," *Econometrica*, Volume 31, July 1963, 464-498.

4. Faddeeva, V., *Computational Methods of Linear Algebra* (C. D. Benster, translator), Dover Publications, New York, 1959.

5. Gale, D., *The Theory of Linear Economic Models*, McGraw-Hill, New York, 1960.

6. Geoffrion, A., "A Parametric Programming Solution to the Vector Maximum Problem, With Applications to Decisions Under Uncertainty," Technical Report No. 11, Graduate School of Business, Stanford University, February 1965. Also issued as Working Paper No. 68, Western Management Science Institute, University of California, Los Angeles. A revised version of the third chapter will appear as "Parametric Concave Programming" in a professional journal.

7. Kuhn, H. and A. Tucker, "Nonlinear Programming," in J. Neyman (ed.), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics*, University of California Press, Berkeley, 1951, 481-492.

8. Theil, H. and C. van de Panne, "Quadratic Programming as an Extension of Classical Quadratic Maximization," Management Science, Volume 8, October 1960, 1-20.